10TH INTERNATIONAL COMMAND AND CONTROL RESEARCH AND TECHNOLOGY SYMPOSIUM
THE FUTURE OF C2

Development of Metrics to Evaluate Effectiveness of Emergency Response Operations

Topic: Assessment, Tools, Metrics
Student Paper

Dr. Donald E. Brown
C. Donald Robinson*
University of Virginia
151 Engineers Way
Charlottesville, VA 22904
Point of Contact:
donaldrobinson@virginia.edu

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **JUN 2005**	2. REPORT TYPE	3. DATES COVERED **00-00-2005 to 00-00-2005**

4. TITLE AND SUBTITLE **Development of Metrics to Evaluate Efectiveness of Emergency Response Operations**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Virginia,151 Engineers Way,Charlottesville,VA,22904**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**49**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Introduction

The overall effectiveness of an emergency response to a large scale disaster is a difficult entity to measure. Every disaster is different and every response is different. In fact, similar events leave totally different results in their aftermath. For example a category two hurricane might inflict much more damage than a category four, depending on the location of impact. Additionally, the emergency response to a flood in rural Mississippi will be totally different than the emergency response to a flood in New Orleans, LA. To effectively measure the overall response to any of these events, proper metrics must be used and interpreted appropriately. This paper will focus primarily on a macro-scale evaluation of the initial assessment and assistance given directly following a natural or manmade disaster. The long term effects and rebuilding of an affected area will not be considered. Furthermore this paper will be constrained to looking at a high level view of assessment rather than at specific methods used to achieve selected goals.

Problem Statement

There are many theories and arguments as to what is the proper or optimal way to respond to an artificial or natural disaster. Should the police help with medical response; should firemen help maintain order; who is in charge and what responsibilities does everyone have? While there are many different methods used, it is difficult to determine which ones are better than others. There are also many different thoughts on where to concentrate efforts and how to use them. While many theories exist, it is very difficult to validate the efficacy of any method due to the inherent irregularity of disasters. No two disasters are alike and thus it is hard to determine whether one method is better than another. Lastly, an actual disaster is not a place to "test" a new idea on response. During

an actual disaster, the response must not be a test of new ideas with inherent confusion, but rather a well tested and trained response.

It is also very cost prohibitive to create an artificial disaster for training and testing. Even a single building collapse scenario involves hundreds of thousands (if not more) of dollars to merely set up the exercise. Often test scenarios like this can only be done only once to test one theory or idea. While the hands-on training for involved personnel is invaluable in a real exercise, the ability to test different methods, in an effort to streamline emergency response to a real disaster, is unavailable. Additionally the people being trained only get one shot to test their skills in a real scenario. Thus, it is difficult to get repetitive training for a particular disaster.

Lastly, relying on historical data from past disasters is necessary and is the basis of any model which might be created. However, using historical data for analysis of an emergency disaster has many pitfalls. As Guha-Sadir and Below explain, this approach of only using historical data for disaster management purposes is reactive and not proactive. There exist many gaps and inconsistencies with historical data due to the "lack of standardized methodologies and definitions" for data measurement during a disaster [8]. Computer simulation would fill some of the holes in real disaster data. Computer models of the available data could be scaled appropriately to represent different scales of disasters. This model could originate from the real analysis of only a few available data sets. Thus, the data from a few disasters, which is difficult and tedious to analyze, could be transformed into many data sets which are more complete and easier to analyze via simulation. Henceforth, using a computer to simulate a large-scale disaster would not

only provide a more controlled environment for data collection but also would generate a more consistent data stream.

Nevertheless, how does one evaluate the output of a computer simulation or interpret a data set from a real disaster. There must be an objective and measurable way to evaluate the efficacy of a particular response. Measuring an emergency response is a difficult task due to the nature of the event itself. The work of Guha-Sadir and Below furthermore address the issue of inconsistent measures. “No internationally standardized method for assessing damage has been put forward for global use” [8]. It is important that any measurement of performance of the overall response should present an accurate depiction of the effectiveness of the responders in completing their mission. This paper will seek to identify such metrics for the purpose of assessing the output of disaster operations in an effort to improve the inputs of the emergency response.

Background

Computer Simulation

Measuring the performance of an entity like an emergency responder is a difficult task. There is no standardized method for obtaining data and information from an actual disaster [8]. Furthermore, relying on data sets of disasters collected from various agencies leads to ambiguous terms, inconsistent and incomplete data, and general confusion [8]. Thus, using real disaster data is tedious and incomplete at best.

Modeling and simulation provides an inexpensive and time-effective method of observing a system and also provides a way to test multiple inputs and evaluate different outputs. As [16] suggests, one of the best reasons for implementing a simulation to observe a system is to, “leverage simulation’s advantages in cost and time relative to

many real-world experiments". The importance of simulation of disasters and emergency response has been noted in [18].

One challenge of integrating micro-scale simulations together to observe a system level progression of a disaster is the inherent disorder present in a disaster scenario. In an emergency response to a disaster the responders must undergo a change in overall strategy from typical day to day response, to disaster response. Disaster response involves more interagency collaboration as well as different tactics and methods of carrying out the overall mission of assistance to the affected population [10]. This change can cause great deals of confusion and must be dealt with in the overall model of a simulation as well as the measurement of an emergency response. Modeling this confusion accurately requires validation of the model. Validation and assessment of a simulation is critical [1] [11]. Validating a simulation of a chaotic event as a disaster is difficult due to the scarcity of data for comparison. Similarly, assessing the validity of selected measurements of a disaster will require verification from various sources.

If a valid computer simulation of a disaster event is created, its outputs must be assessed. As Jain and McLean state regarding the challenges of a disaster simulation, "Interpretation of the simulation output data might not be a straightforward process. . . . it may not be clear what action should be taken based on the results" [5]. This seems to be a common problem in simulation of disasters and other events. Considerable effort has been expended in creating simulations. However, often the output is not clearly addressed. As stated in [5], the objective of response to a disaster is to "minimize the impact of disaster events on entities of interest". The specific metrics which describe the

“impact” are not identified and are often just assumed. For true objective analysis of any simulation or actual response, a clear and defined method should be consistently used.

Describing a Disaster

To measure a response to a disaster, the terms response and disaster must be clearly defined. As [6] writes, disasters are spatial-temporal events which impact social units which then invoke responses to the event or events. Furthermore, [12] goes on to write that “Responses may involve structural engineering (physical), be relevant before, during, or after the impact is felt (temporal), or result from a variety of social and organizational processes (social)”. Thus responses to disasters are essentially an effort to restore the previous status quo to an affected area. Thus, response to a disaster will be defined simply as *turning chaos into order*.

In [10], the author addresses some of the pertinent issues specific to emergency medical response to a large scale disaster. Such factors include the “golden hour” which specifies that a casualty must be identified, stabilized and transported to a hospital in one hour to maximize the likelihood of survival. This is one standard which EMS personnel use in normal operations. However, in a large-scale disaster this standard often cannot be met due to strains on resources available for treatment. The paper then describes how EMS systems have to adapt to accomplish such standards and how to model such adaptations via chaos theory. Thus time is a critical element in assessing the effectiveness of many aspects of emergency response.

Furthermore, the Federal Emergency Management Agency has specified 4 general stages of emergency management: (1) preparatory or planning period including policies and programs for impact mitigation; (2) preparedness or training and positioning of

resources; (3) immediate response to a disaster including assessment, resource allocation and command and control; (4) and recovery after the disaster [15]. Furthermore in [20], the authors compiled a list of informational requirements for each stage of a disaster.

Effectiveness Measurement

Assessing the effectiveness of an organization is a difficult task. As written in [4], “Although the literature on organizational effectiveness is large and growing, there seems to be little consensus on how to conceptualize, measure, and explain effectiveness”. The author continues to address multiple ways of achieving effectiveness measures. The three main models of effectiveness are the “behavioral-attitudinal” model, the “processual” model, and the “goal attainment” model [4][3][17]. The author in [4] chooses the goal attainment model to address fire protection performance. Furthermore he explains variance in performance in terms of internal organizational processes and environmental variables.

Another example of using a “goal attainment” approach to measuring disaster management effectiveness was done by [9]. In this article the author suggests a client-stakeholder relationship in measuring disaster managers’ effectiveness for Israel’s Home Front Command. The author compares the stated goals of the organization and the “perception of their provision” as seen by the stakeholders in the system. Analysis was done on the factors influencing the delivery of provisions as well as the factors influencing the perception of service provided. This article lends more evidence for the validity of a goal attainment model for measuring organizational effectiveness in the realm of disaster response.

Additionally, there are several methods of implementing a goal-attainment model for assessing effectiveness. The different models exist due to the difficulty in defining what the goals of an organization should be. [14] explains the "prescribed" method in which the upper management of an organization resembles the true nature of an organization and thus is an appropriate source of goal definitions. [21] discusses the "derived" approach to a goal attainment model of effectiveness assessment. In such a model, the goals are derived from the theoretical function of the agency in question. Thus, the measures of effectiveness are independent of the methods used to achieve the goals.

Also, in an assessment done by the United States Coast Guard [19], a "balanced scorecard" method was created for evaluating the effectiveness of response primarily to an event such as the Exxon-Valdez oil spill. The methodology in obtaining the metrics presented is basically a goal-attainment approach. This model will serve as a reference in obtaining more generalized metrics, not just for a pollution incident.

The differences in measuring effectiveness and measuring performance should also be noted. As stated in [22] regarding the MORS approach to measures of effectiveness, measures of performance (MOP's) are "related to inherent parameters (physical and structural) but measure attributes of system behavior". Whereas measures of effectiveness (MOE's), in a C2 context, "Measure how the C2 system performs its functions within an operational environment". [23] earlier had defined MOE's as "The measure of effectiveness is the criterion by which solutions will be judged – proposed solutions, solutions under test, or solutions in being". This definition, while not necessarily representing the most current research in MOE's, represents, very simply

what this paper is trying to achieve. MOP's are measures depicting how well an entity performed in accomplishing a mission, whereas MOE's depict how well the mission was accomplished without looking at the method in which it was accomplished.

It is therefore appropriate to use MOE's in designing a method to measure how well an emergency response entity accomplishes its missions in the wake of a disaster. Again, the purpose of this paper is to develop metrics which will be used to evaluate and test different methods and technologies used in emergency response in an effort to determine which ones have the most beneficial outcome for an affected population. Thus, it does not necessarily matter *how* the mission was accomplished, rather *how well* it was accomplished. This is of course evaluating it from a macro scale point of view. As one probes deeper into the details of an emergency response, MOP's become necessary and appropriate. However, for the purposes of this paper, only the high level view of a response will be considered.

There are also three issues which need addressing when presenting MOE's. These issues were addressed by [24] and are: "(1) aggregating the measures, (2) ensuring that the measurements have consistent units and a 'direction of improvement', that is, measurements need to be formulated so improvements add to the overall improvement of the system; and (3) being able to combine quantitative and qualitative measures[25]".

A Framework

To develop the MOE's necessary, a framework in which to describe the measures must be explicitly stated. Past research suggests that a goal-attainment approach seems not only appropriate but also best for these metrics. A goal attainment approach is not only scalable, in that goals can be broadly defined to be applicable on many scales, but

also dynamic. As new goals are defined, new metrics can be easily added to address the goals. This can be done without necessarily having to measure the other metrics again. Such would not be the case in a survey of people for a behavioral effectiveness model. Also, the derived approach seems appropriate here in that this paper seeks to define metrics to use based on theoretical functions. How that function is implemented is not really the point. These metrics are designed so that many different ways to accomplish the goals of emergency responders can be evaluated. Having the measure of effectiveness dependent on the method used would be counterproductive.

The first element of the framework within which these metrics are developed is defining disaster. A disaster can be anything from the death of a parent to a widespread outbreak of influenza. Disasters also can range from a stock market crash to a hurricane. For the purposes of this paper a disaster will be defined in the context of the commonly known natural or manmade disasters. Thus, a disaster will be defined as any spatial-temporal event which negatively affects an underlying community's population and/or property. Furthermore, the disaster will be characterized as being large enough that the resources the community has to mitigate the disaster are stretched beyond the limits of their capacity. Such events that a community can readily cope with, such as small fires, individual medical incidents and car crashes, while tragic to any victim, will not be considered a disaster in this sense.

An emergency responder needs to also be clearly defined. In the context of this paper and emergency responder will be considered any person who actively engages in organized assistance and mitigation efforts to restore the underlying community to its state prior to the disaster. While most people consider people like police officers and fire

department personnel to be the emergency responders, one must not overlook the efforts of people like ham radio operators and helicopter pilots when looking at responders in a disaster.

Also, as mentioned before, the Federal Emergency Management Agency has defined four stages of a disaster [15]. The first two stages really deal with the positioning of human capital and public resources to prepare for a disaster. The last stage addresses the actual rebuilding of a community after the disaster's immediate effects have been mitigated. This paper will only address the evaluation of response efforts in the third stage of a disaster. This is the immediate assistance phase where emergency responders are typically recognized as having the greatest impact on the population.

Lastly, this paper will deal with the evaluation of response efforts in this initial assistance phase of a disaster and will do so from a high-level point of view. A high-level, or system level view is used because of the inherent diversity of disasters. Some disasters might only damage property in a geographic area, like a hurricane, while another disaster might only due damage to humans, like a disease outbreak. Regardless of the type of disaster, emergency responders must always be prepared for anything. For example, while the probability of an earthquake in Virginia is low compared to California, the emergency responders in Virginia must have some level of preparation for such an event. This is because when it happens, they will be the ones called on the assist the underlying population. Thus, one must look at the very general goals when trying to evaluate the effectiveness of an entire response system as it responds to any type of disaster.

Measuring a Disaster

Any measurement of a response must be taken in context of the disaster itself. This is necessary if one wants to compare responses to different disasters. However, measuring the disaster itself is equally as difficult as measuring the response. Hence more attention will be given to the response in this paper, but the measurement of the disaster will be addressed. The results shown here are a compilation of expert interviews, literature research, and personal experience as an emergency responder.

The two types of disasters that are primarily thought of are natural and manmade. Natural disasters typically affect a larger geographic region than manmade disasters. However, manmade disasters often affect more densely populated regions that natural disasters. Thus, they are inherently different is scope implying several aspects need to be observed. However, the primary characteristics of a disaster are the people and property affected.

The first metric for quantifying a disaster is the area of the region affected by the disaster. This can be measured as a two-dimensional measure of geographic space. This gives an overall sense of the scope of the effects.

A measure of the effects to the lives of the population is also necessary. This can be measured as the percentage of the population in the affected region that sustains injury. Representing the injuries as a percentage of the potentially affected population gives a natural normalization. This facilitates direct comparison between dissimilar events. Furthermore, the injuries should be broken down in to types of injuries as well as severities. For instance, the percentage of casualties in a region could be represented by an “m by n” matrix where m is the total number of types of injuries and n is the total

number of levels of severities. This allows a researcher to easily see the breakdown of all injuries relative to the entire population with little effort. An example is as follows:

	Mild	Moderate	Severe	Death	
Respiratory	2.8%	1.9%	0.3%	0.2%	5.2%
Cardiac	2.5%	1.8%	0.3%	0.2%	4.8%
Broken Bones	11.0%	7.6%	1.3%	0.9%	20.8%
Crushing	5.9%	4.1%	0.7%	0.5%	11.2%
Lacerations	24.5%	16.8%	2.9%	1.9%	46.1%
Burns	2.0%	1.4%	0.2%	0.2%	3.8%
Unknown	4.3%	3.0%	0.5%	0.3%	8.1%
	53.1%	36.5%	6.3%	4.1%	

The property damage in an area must also be quantified. This should be done in a similar fashion. Representing the property damage as the percentage of dollar loss with respect to the original dollar value gives a natural normalization. Also, by breaking the property damage down in a similar fashion as the casualties gives a researcher more information with which to base results. Thus the damage can also be represented as an “m by n” matrix where m is the number of property types that could sustain damage and n is the number of levels of damage. Each entry in the matrix is the percentage of each certain type of property that sustained a particular level of damage as a result of the disaster. An example is as follows:

	Mild	Moderate	Severe	Total	
Concrete Building	6.3%	1.2%	0.2%	0.1%	7.9%
Wood Building	11.9%	2.3%	0.4%	0.3%	14.8%
Metal Building	16.7%	3.2%	0.6%	0.4%	20.8%
Roadway	9.0%	1.7%	0.3%	0.2%	11.2%
Bridge	8.1%	1.6%	0.3%	0.2%	10.1%
Transportation Vessel	9.8%	1.9%	0.3%	0.2%	12.2%
Power Infrastructure	11.3%	2.2%	0.4%	0.2%	14.1%
Communications	7.1%	1.4%	0.2%	0.2%	8.9%
	80.1%	15.4%	2.8%	1.7%	

In this example, the row for roadways can be interpreted as 9% of all roadways sustained mild damage which may be characterized as damage amounting to between 10% and 25% of the total replacement value of the road. Also 1.7% of all roadways sustained damage amounting to a value of between 25% and 75% of the total replacement value of the roadway and .3% of all roadways sustained damage amounting to between 75% and 90% of the total replacement value for the roadway. Lastly, .2% of roadways sustained total damage meaning total replacement of the road. The last percentage value is interpreted as 11.2% of all roadways sustained damage.

Measuring the Response

Scope of Response

Response to a disaster must be appropriately scaled according to the size of the disaster. Because response, for this paper, is defined in terms of an organized response, it will be the extent of this organization which will reflect the size of the disaster response. Currently the United States is undergoing an effort to standardize the method in which disasters are managed. It is called the National Incident Management System [25]. This system standardized an Incident Command System which classifies each individual responder into a particular sector. Within the system there might be different sectors with the same purpose, but it is designed so that the "span of control for any one person never exceeds 5-7 people" [27]. The system is designed to scale appropriately to handle any disaster regardless of the size. As stated by [27], "The incident command system is set up to break even the largest of incidents into manageable pieces". Thus, by measuring the number of branches and sectors initialized under the National Incident Management System guidelines, the entire scope of the rescue effort could be represented in a

standardized manner, regardless of the type of disaster. However, this implies that such a system would be used in the response efforts of the disaster. Because a disaster is defined as an event which is large enough to exceed the capabilities of a community to handle the effects, it can be assumed that some system of organization has to be implemented to handle new agencies assisting in response. The NIMS is the only nationally standardized method of emergency management, outside of military protocol, and thus should be the benchmark for measuring the extent of the response.

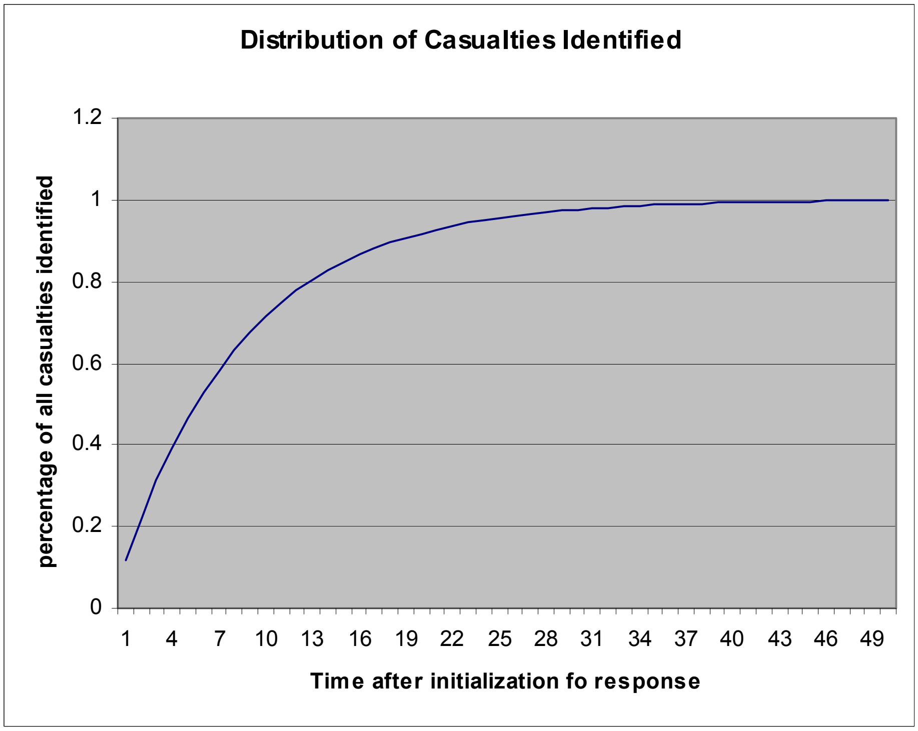
Goals of Response

Next the goals of the actual response must be identified. Several expert interviews as well as personal experience have led to the conclusion that emergency responders' goals basically fall into two categories, protect life and protect property [27][28][29][30]. There are also several aspects to protecting life and property during the initial phases of a disaster. These categories basically are stabilization, rescue, mitigation, and safety. Stabilization involves stabilizing any obvious threats to the population or property to the greatest extent possible so as to contain them. Rescue involves searching for and identifying casualties and properly triaging them prior to transport to a hospital or other staging facility. Mitigation involves preventing further harm to a population by diminishing the malicious effects of a disaster and restoring basic services. Mitigation goes hand in hand with stabilization in some respects. For example, in a flood situation, stabilization would involve placement of sandbags to divert water away from populated areas while mitigation would involve evacuating the residents from the potentially affected area. Lastly safety involves ensuring that the responders accomplish their mission effectively by not becoming incapacitated. When a responder is injured, not only

are more victims put at risk, but also that responder then becomes yet another victim that needs aid. Thus it is imperative that in any response, those who are assisting others do so safely and in a manner so as to be able to assist others throughout the duration of the incident.

The Metrics

Thus, these goals provide the basis of a goal-oriented approach to defining proper system-level measures of effectiveness for response to a disaster. The first metric which seems almost paramount to any other would involve the saving of lives. It is generally agreed in the United States that people are more important than property. Thus the first metric would be the total distribution of casualties with respect to the time they are found. This metric would represent how well the responders search and identify those who need help. A total distribution is used for it contains more information than just a singular statistic. It is clear that the optimal situation is one where all of the casualties of a disaster are identified within a short time frame. A graph of this measurement would look similar to a random variable's cumulative distribution function. An example is shown here:



While during the actual course of responding to the disaster it is impossible to know how many total casualties there are, this data could be later calculated for comparison purposes. Furthermore, this distribution could be analyzed for each type of injury or each severity. A probable case might be that the majority of mild injuries would be found in a short time period whereas it would take longer to identify all of the deaths associated with the disaster. Also, *identification* is defined as being recognized by some level of command of the disaster. This level of command could be a dispatch center, crew chief, or any other commander. Thus, identification is really defined by the ability of responders to be mobilized to assist that victim. This mobilization is initiated from some command level and thus that commander's knowledge of the victim is defined as identification.

This type of distribution could also be measured for the amount of property damage in the area. Assessing the condition of the infrastructure is critical to successful

operations because planners need to know the state of the infrastructure so as to implement the most efficient response. Assessing damage also will coincide with assessing casualties because many casualties are caused by property damage. Thus a distribution like the one for casualties should be used as a measure of assessing total damage.

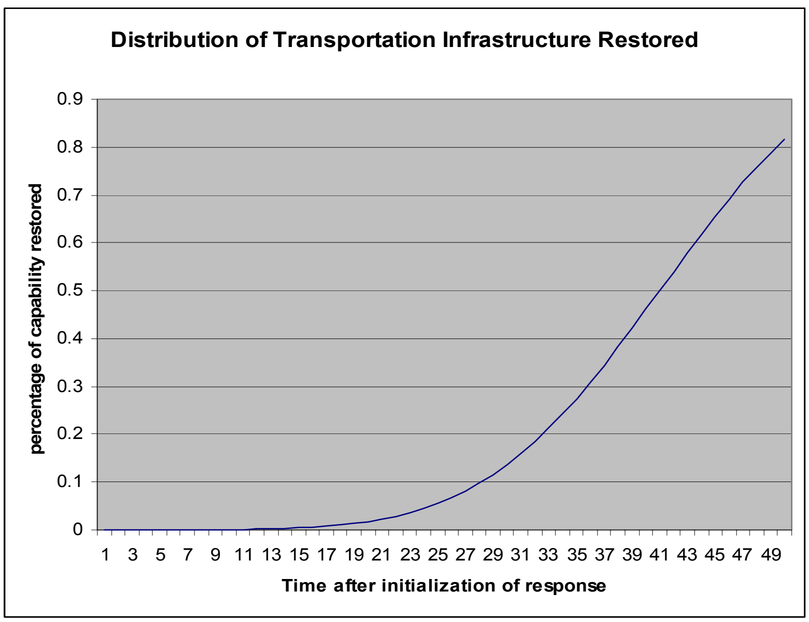
The next metric for measuring the effectiveness of a response is how well the responders deal with the casualties that are found. This is where standards like the "golden hour" are currently used. Basically, it does no good for a responder to identify a victim without being able to give any aid to that victim. Aid consists of triage, stabilization, and then medical intervention. However, all of these can be simply measured by looking at the percentage of victims whose condition worsens after being identified by emergency responders. This can be further broken down into what severity levels the casualties transitioned to from the identified state. There are some types of incidents where this metric would not necessarily represent the efforts of responders to aid casualties. An example would be a radiological contamination. There might be a large number of victims who are identified, yet responders might not be able to give any aid due to the radioactive contamination in the area. Thus a large percentage of those victims' conditions would potentially get worse even after being identified. However, this is not necessarily due to the ineffectiveness of the responders. This metric must therefore be taken in context of the disaster. However, in a general sense this type of measurement would give a general idea of how effective the response was in administering aid to the victims as well as measuring the effectiveness of rescue operations.

A similar metric to the one mentioned above represents the responders' efforts to save property. It is thus the percentage of property that sustained further damage after the onset of disaster response. Once again this would be measured as a percentage of the total damage. In other words the metric might be interpreted as "of the damage in the region, 32% of it either occurred or got worse after the onset of organized response". This helps assess not only how well the responders save property, but really how well the scene is stabilized. Stabilization implies a steady state. However, in a disaster it seems as though nothing is steady. Thus, to measure this stability, one should only look at how much of the scene gets worse. It is the job of those involved in the recovery effort to make everything "better". The initial responders are really there to just stabilize the area and this is measured by the percentage of property that sustained further damage after the onset of disaster operations.

Another important metric for measuring stabilization is the time from the onset of the disaster until the time when the scene is declared by the top commander as being "under control". Such a determination represents a shift from assessment and stabilization, to completing missions and moving towards recovery. While such a determination does not mean that rescue or initial response operations are to be ended, it simply implies that command has done a complete assessment, implemented a plan, and has a clear picture of future needs for the response. In a disaster response operation it is desirable to make this metric as small as possible. Emergency managers need to stabilize the scene quickly so that unencumbered rescue and recovery operations can proceed.

Along with this metric would be a distribution of the time until a certain percentage of a community's total functionality was restored. Here, functionality can be

defined in terms of both infrastructure capability as well as human capabilities. An example would be the distribution of time until a certain percentage of roadways have been restored to a level of operability. Thus this metric would be divided into two several distributions. Examples of the functionality distributions are ones for communications infrastructure, transportation infrastructure, medical infrastructure, sanitation infrastructure, and others. Each point in the distribution would represent the total functionality of the system in question at a time after the initialization of organized response. An example follows:



This distribution will not necessarily reach a value of 1. The last point in the distribution would really be up to the discretion of the researcher. This is because there is not necessarily a definite time when initial response operations end and recovery operations begin. However, theoretically, it is this point which should be the end of the distribution. At this time a lot of the capabilities in question might not be fully restored. Thus, at the end of initial response operations, only 30% of full transportation capabilities might be

restored; the distribution of this would represent the history of it being restored from the initial state to that 30% level.

The last element in a disaster response is to measure the safety of the response with respect to the responders. Safety is paramount in any response. Responders are typically trained in some field well enough that one responder potentially could assist multiple casualties. Therefore if one responder is incapacitated, many more victims of the disaster are put in danger of not receiving adequate care. Thus, in any disaster scenario one must measure how well the responders take care of other victims as well as how well the responders assist themselves. Thus the final metric would be a breakdown of responder injuries and severities. This would be done in a similar fashion to that of casualty measurement for measuring the disaster itself. The responder casualty breakdown would be the same "m by n" matrix as that of the casualties where m represents the number of types of injuries and n represents the levels of severity.

A Note on Hierarchy and Subjectivity

The overall emergency response to a disaster has a very hierarchical nature. There is a hierarchy of organization, of goals, and of priorities. The metrics discussed here are a basis for a high-level view of the overall response. There are many other attributes to fully measuring every aspect of response. Measurements like gallons of water pumped at a fire or amount of medical supplies used to assist victims, while important, do not represent a total picture of the overall response. These types of measures fall into a subcategory of response and thus could be considered, but at a lower level.

Furthermore, these types of measure really address MOP's rather than MOE's. As mentioned before, MOP's are really measures of how well an entity performed in accomplishing a mission. MOE's really address how well the mission was carried out regardless of the method in which it was accomplished. Thus, most any other type of measurement should fall under the general goals of protecting life and property. Furthermore the general goals of stabilization, mitigation, rescue and safety

There are also subjective measures of effectiveness which are important, though not addressed. Measure like the public's perception of the response, while important to policy makers and emergency response leadership, is a subjective measure not objectively quantified. Some of the metrics presented here might have some subjectivity as to their implementation. For example, whether someone is classified as a mild or severe casualty might be a subjective determination. However, the fact that the person is a casualty and is injured to some level is objective. Thus the metrics are meant to be directly measurable and used not only in a real world disaster, but also for a computer simulation.

Summary

In conclusion, a derived goal attainment approach was used to define primary goals of emergency responders to a disaster and then to define metrics to quantify the effectiveness of the responders in accomplishing those goals. Disaster in this paper was defined as a spatial-temporal event which negatively affects some population beyond its ability to respond to the event. A disaster could be measured with the following metrics.

- Total area of affected region
- Matrix of percentage of casualty severities with respect to injury types

- Matrix of percentage of property damage by value using severity levels with respect to property types

A responder to a disaster is defined as any person who actively engages in an organized effort to assist victims and mitigate the effects of a disaster. The main goals for responding to a disaster are first and foremost, the protection of people and property. To do this the main goals are stabilization, mitigation, rescue, and safety. The metrics defined to measure the effectiveness of responders in carrying out these missions are as follows.

- The total number of branches and sectors of the initiated command structure most likely defined by the National Incident Management System
- Distribution of percentage of casualties identified versus the time after the initialization of response efforts
- Distribution of percentage of total property damage identified versus the time after the initialization of response efforts
- Percentage of casualties of each injury type whose condition worsens after being identified by responders
- Percentage of property of each property type that sustains further damage after being identified by responders
- Time until the disaster region is declared under control by the incident commander
- Distribution of percentage of a community's human capabilities and infrastructure functionality versus time after the initialization of response efforts
- Matrix of responder casualty severities with respect to injury types

These metrics objectively assess the level at which responders accomplish the aforementioned goals. Furthermore, these metrics are designed to facilitate direct comparison of different responses to different disasters. The metrics also scale appropriately for different disasters and have clear “direction of improvement” thus allowing good comparison of events.

References

[1] Balci, O. 1998. Verification, Validation, and Testing. *Handbook of Simulation: Prinicples, Methodology, Advances, Applications, and Practice*. New York, NY: John Wiley and Sons.

[2] Cameron, Kim S. 1986. A Study of Organizational Effectiveness and its Predictors. *Management Science*. Vol. 32, No. 1. pp. 87-112. Pub. INFORMS.

[3] Campbell, John P. 1977. On the Nature of Effectiveness. *New Perspectives on Organizational Effectiveness*. Ed Paul Goodman and Johannes Pennings. Pp.13-55, San Francisco: Jossey-Bass

[4] Coulter, Philip B. 1979. Organizational Effectiveness in the Public Sector: The Example of Municipal Fire Protection. *Administrative Science Quarterly*, Vol. 24, No.1, pp.65-81. Pub. Johnson Graduate School of Management.

[5] Jain, Sanjay and Charles McLean. A Framework for Modeling and Simulation for Emergency Response. *Proceedings of the 2003 Winter Simulation Conference.* pp. 1068-1076, 2003.

[6] Fritz, C.E. 1961. Disasters. In *Social Problems*, ed. R. Merton, R. Nisbet, pp.651-94. New York: Harcourt, Brace and World.

[7] Goodman, Paul S., Robert S. Atkin and F. David Schoorman. 1983. On the Demise of Organizational Effectiveness Studies. *Organizational Effectiveness: A Comparison of Multiple Models*. Ed. Kim Cameron and David Whetton. P.163-183. New York: Academic Press.

[8] Guha-Sapir, Debarati and Regina Below. 2000. *The Quality and Accuracy of Disaster Data: A Comparative Analysis of Three Global Data Sets*. A study by the Provention Consortium.

[9] Kirschenbaum, Alan. 1994. Measuring the Effectiveness of Disaster Management Organizations. *International Journal of Mass Emergencies and Disasters,* Vol.

22, No. 1 pp. 75-102. Pub. Chool of Policy, Planning, and Development, Univ. Southern California.

[10] Koehler, Gus. 1995. What Disaster Response Management can learn from Chaos Theory. Ed. Gus A. Koehler. *What Disaster Response Management Can Learn From Chaos Theory,* May 1995, California. California Research Bureau.

[11] Knepell, P.L. and D.C. Arangno. 1993. *Simulation Validation: A Confidence Assessment Methodology*. IEEE Computer Society Press.

[12] Kreps, G.A. 1984. Sociological Inquiry and Disaster Research. In *Annual Review of Sociology*, Vol. 10 pp.309-330. Pub. Annual Reviews.

[13] Parks, Roger B. 1984. Linking Objective and Subjective Measures of Performance. *Public Administration Review*. Vol. 44, No.2 pp.118-127. Pub. American Society for Public Administration.

[14] Perrow, Charles. 1968. Organization Goals. *International Encyclopedia of the Social Sciences*, Vol. 11:pp. 305-311. ed. David Sills. New York: MacMillan Free Press.

[15] Petak, W. J. and A. Atkisson. 1982. *Natural Hazard Risk Assessment and Public Policy*. New York: Springer-Verlag.

[16] Sanchez, Susan M. and Thomas W. Lucas. Exploring the World of Agent-Based Simulations: Simple Models, Complex Analyses. *Proceedings of the 2002 Winter Simulation Conference*. pp.116-126, 2002.

[17] Scott, William R. 1977. Effectiveness of Organizational Effectiveness Studies. *New Perspectives on Organizational Effectiveness*. Ed Paul Goodman and Johannes Pennings. Pp.63-95, San Francisco: Jossey-Bass

[18] Sullivan, Thomas J. Modeling and Simulation for Emergency Response. Lawrence Livermore National Laboratory, Report No. UCRL 92001 Preprint.

[19] United States Coast Guard Office of Response. "Measuring Response: A Balanced Response Scorecard for Evaluating Success." www.uscg.mil/hg/g-m/mor/articles/proceed.pdf

[20] Wallace, W. A. and F. De Balogh. 1985. "Decision Support Systems for Disaster Management". *Public Administration Review,* Vol. 45, Special Issue: Emergency Management: A Challenge for Public Administration, pp. 134-146. Pub. American Society for Public Administration.

[21] Yuchtman, Ephraim and Stanley Seashore. 1967. A System Resource Approach to Organizational Effectiveness. *American Sociological Review*, Vol. 32, pp.891-903.

[22] Green, John M. and B.W. Johnson. 2002. "Towards a Theory of Measures of Effectiveness." Presented at CCRTS, June, 2002.

[23] Goode Harry H. and Robert E. Machol. 1957. *System Engineering; An Introduction to the Design of Large-Scale Systems*. NewYork: McGraw-Hill.

[24] Smith, Neill and Thea Clark. 2004. "An Exploration of C2 Effectiveness – A Holistic Approach". Presented at CCRTS.

[25] Filar, Jerzy, L. C. McArthur, Y. Yue and M-L Wu. 2001. "Integration of Measures of Merit in the Context of Network Centric Warfare", DSTO Draft Report.

[26] Federal Emergency Management Agency, National Incident Management System. http://www.fema.gov/nims/, accessed Mar 10, 2005.

[27] Boyd, Carr. Interview on Mar 1, 2005.

[28] Cannon, Robert. Interview on Mar 8, 2005.

[29] Marks, Craig Allen. Interview on Mar 9, 2005.

[30] Townsend Kenneth. Interview on Mar 2, 2005.

Development of Metrics to Evaluate Effectiveness of Emergency Response Operations

10th International Command and Control Research and Technology Symposium

The Future of C2

Dr. Donald E. Brown
C. Donald Robinson*

Executive Summary

- 3 metrics are developed to describe the disaster event itself
- 8 metrics are developed to measure the effectiveness of response
- Primarily to be used in computer simulation with total information
- Could be used in a real world scenario

Primary Motivation

- There will always be another disaster
- We can always do better in our response
- Need a method of comparison (Better than What!)

Photo Courtesy of FEMA [2]

Problem Statement

- Idea of total preparedness
- Decentralized system of response
- Impossible to treat every locality individually
- How can we assess our ability to respond to catastrophic events

Why?

- Lack a standardized system of measurement necessary for comparison [1]
- Need an objective way to identify deficiencies in response in order to improve
- Global applicability
- Funding

Define Disaster

- Spatial-temporal event which abnormally negatively affects some population beyond its ability to mitigate the effects of the event
- Natural
- Artificial

Define Responder

- Any person who actively engages in an organized effort to assist victims and mitigate the effects of a disaster
- Protection of lives and property

The Use of Simulation

- Using real data is reactive and not proactive [1]
- Simulation provides a safe, cost-effective way to train responders and test technology
- Provides perfect information for more precise measurements

Measuring a Disaster

- Geographic Scope
- Scope of Injuries
- Scope of Property Loss

Disaster Metrics

- Total 2-D area of the region primarily and most directly affected by the disaster
- Matrix of injury types and severities as a percentage of the total population affected
- Matrix of property damage types and severities as a percentage of the total property affected

Example of Injury Matrix

	Mild	Moderate	Severe	Death	
Respiratory	2.8%	1.9%	0.3%	0.2%	5.2%
Cardiac	2.5%	1.8%	0.3%	0.2%	4.8%
Broken Bones	11.0%	7.6%	1.3%	0.9%	20.8%
Crushing	5.9%	4.1%	0.7%	0.5%	11.2%
Lacerations	24.5%	16.8%	2.9%	1.9%	46.1%
Burns	2.0%	1.4%	0.2%	0.2%	3.8%
Unknown	4.3%	3.0%	0.5%	0.3%	8.1%
	53.1%	36.5%	6.3%	4.1%	

Example of Damage Matrix

	Mild	Moderate	Severe	Total	
Concrete Building	6.3%	1.2%	0.2%	0.1%	7.9%
Wood Building	11.9%	2.3%	0.4%	0.3%	14.8%
Metal Building	16.7%	3.2%	0.6%	0.4%	20.8%
Roadway	9.0%	1.7%	0.3%	0.2%	11.2%
Bridge	8.1%	1.6%	0.3%	0.2%	10.1%
Transportation Vessel	9.8%	1.9%	0.3%	0.2%	12.2%
Power Infrastructure	11.3%	2.2%	0.4%	0.2%	14.1%
Communications	7.1%	1.4%	0.2%	0.2%	8.9%
	80.1%	15.4%	2.8%	1.7%	

How to Measure Effectiveness

- Measuring Effectiveness vs. Performance [3][4]
- The Goal-Attainment approach
- Different ways to accomplish the same goal
- Measure Effectiveness regardless of disaster type

Goals

- Turn Chaos into Order
- Protect Lives and Property
- Stabilization
- Rescue
- Mitigation
- Safety

Constraints

- FEMA's 4 stages of a disaster
- Only dealing with immediate assistance phase
- Only measuring the effect of *organized* response

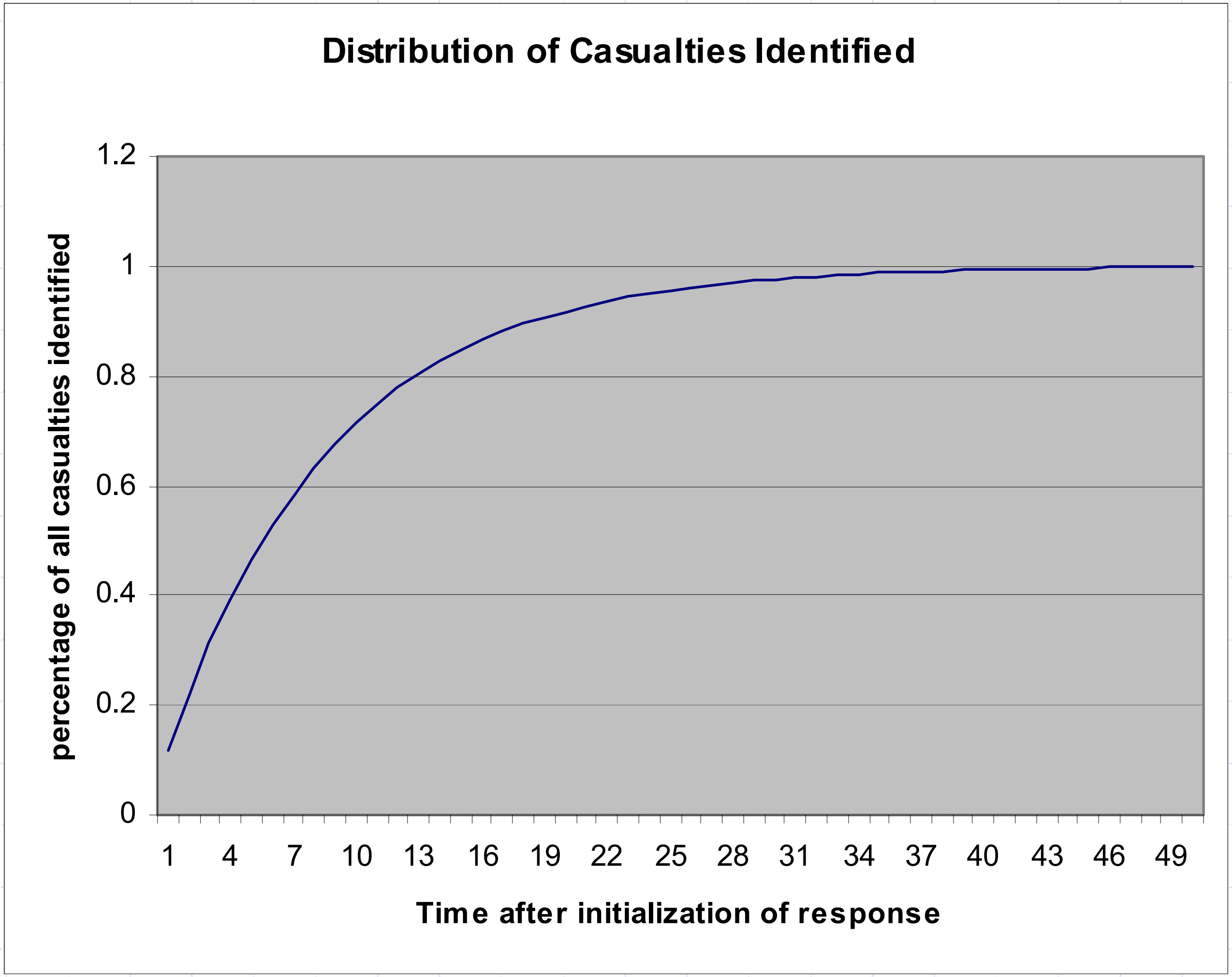
Response metrics (Scope and Safety of the Response)

- Total number of branches and sectors of the initiated command structure as defined in the National Incident Management System
- Time from initialization of response until the disaster region is declared under control by the incident commander
- Matrix of responder casualty severities with respect to injury types

Response Metrics(Rescue)

- Distribution of percentage of casualties identified versus the time after the initialization of response efforts
- Distribution of percentage of al property damage identified versus the time after the initialization of response efforts

Example of Distribution of Identified Casualties



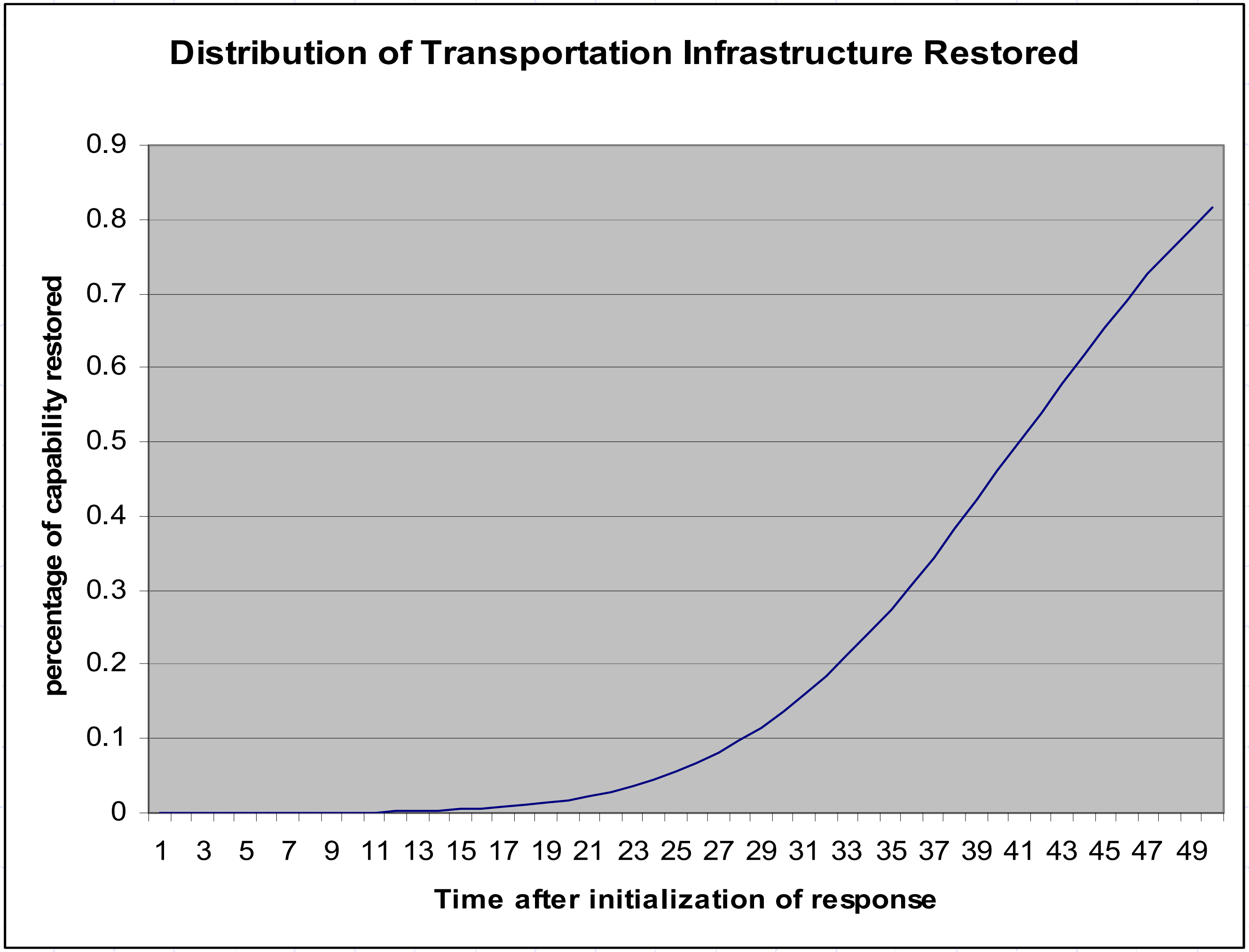
Response Metrics (Stabilization, Rescue, Mitigation)

- Percentage of casualties of each injury type whose condition worsens after being identified by responders
- Percentage of property of each property type that sustains further damage after being identified by responders
- Also the same percentages except for those casualties and damage that are not identified by responders

Response Metrics (Stabilization, Mitigation)

- Distribution of percentage of a community's human capabilities and infrastructure functionality versus time after the onset of response efforts

Example of Distribution of Restoration of Infrastructure



Final Comments

- Take disaster type into consideration only after metrics have been assessed
- Interpret metrics in context
- Future work: refining categories like injury types and severities
- Future work: proving statistical sufficiency of metrics

References

- [1] Guha-Sadir, Debarati and Regina Below. 2000. *The Quality and Accuracy of Disaster Data: A Comparative Analysis of Three Global Data Sets*. A study by the Provention Consortium.
- [2] Federal Emergency Management Agency, Photo Library. http://www.photolibrary.fema.gov/.
- [3] Green, John M. and B.W. Johnson. 2002. “Towards a Theory of Measures of Effectiveness.” Presented at CCRTS, June, 2002.
- [4] Goode Harry H. and Robert E. Machol. 1957. *System Engineering; An Introduction to the Design of Large-Scale Systems*. NewYork: McGraw-Hill.